Sparse Locally Adaptive Cost Aggregation For Stereo Matching

Jinjin Li
Arizona State University
jinjinli@asu.edu

Lina Karam
Arizona State University
karam@asu.edu

Abstract

Local Stereo Matching methods have been attracting more attention recently because they are faster and more suitable for real-time applications. In this paper, we propose a novel Sparse Locally Adaptive Cost aggregation (SLAC) local stereo matching method in order to improve the disparity accuracy while preserving fast computation speed. The proposed method consists of a fast initial cost aggregation stage followed by a refined cost aggregation that is only performed over a sparse subset of disparities. The cost aggregation is performed in a locally adaptive manner by adapting the support region to the local image intensity and structure. A novel sparse disparity subset selection method is proposed by assigning a significance status to candidate disparity values, and selecting the significant disparity values adaptively. Experimental results tested on Middlebury image sets show better disparity accuracy of the proposed method as compared to other recent local stereo methods.

1. Introduction

Disparity estimation using a stereo image pair has been thoroughly studied in computer vision. According to the survey in [7], there are two broad categories of stereo matching algorithms: global methods [11, 8, 5] and local methods [2, 4, 15, 9, 3]. Local stereo matching methods are much faster and more suitable for real-time applications. Recent local stereo matching methods [12, 2, 4, 15, 9, 10, 3] generate more accurate disparity maps that outperform many global optimization based algorithms. Thus local stereo matching methods have been attracting more attention recently.

In local methods, the cost aggregation steps can affect significantly the disparity accuracy. Recent local methods perform cost aggregation using locally adaptive filtering techniques with region of support and weights that adapt to the local image characteristics. Using a guided filter for stereo matching was adopted in several stereo matching approaches [4, 15, 9, 10, 3]. The guided filter is more suitable for real-time applications due to the linear computation and the speed-up implementation using integral image methods. The adaptive support window methods can also improve the cost aggregation accuracy, thus benefiting the disparity estimation. An efficient adaptive support window approach called the cross-based support region method was firstly proposed in [13], and has been adopted in several recent stereo matching methods [11, 15, 16]. This method can handle the variant shapes near depth discontinuities more accurately and its cost aggregation is more efficient as compared to other adaptive support window methods.

In local stereo matching methods, it is important to reduce the computational complexity of cost aggregation for real-time purposes, while retaining the accuracy of the estimated disparities. There are several local stereo matching schemes trying to reduce the computational load and the aggregation redundancy [1, 14]. The method proposed in [6] presents a fast box-filtered cost volume computation to reduce the redundancy using a disparity subset determined jointly from the disparity histogram and a sampling technique in a square support region. But the fast box-filtered cost volume computation of [6] is noisy and inaccurate, since it makes use of a fixed squared support region that does not adapt to the local image characteristics. Another drawback of the method in [6] is that it adopts a bilateral adaptive weight calculation with a relatively high computational cost.

2. Contribution and Experimental Results

We propose a novel Sparse Locally Adaptive Cost aggregation (SLAC) local stereo matching method. The proposed SLAC local stereo matching method consists of a fast initial cost aggregation stage followed by a refined cost aggregation that is only performed over a sparse subset of disparities. In the proposed method, the support region is calculated in a locally adaptive manner by adapting the support region to the local image intensity and structure. We propose a locally adaptive cross-based support region calculation method that incorporates the variance of the local color change at each pixel in the color similarity threshold. In order to reduce outlier disparity values that correspond to mis-

Algorithm	Tsukuba			Venus		
	Nonocc	All	Disc	Nonocc	All	Disc
Proposed	**1.12**	**1.60**	**6.03**	**0.16**	**0.42**	**1.71**
AdaptiveGF	1.04	1.53	5.62	0.17	0.41	1.98
HistAggr2	1.93	2.30	6.39	0.16	0.46	2.22
CrossLMF	2.46	2.78	6.26	0.27	0.38	2.15

Algorithm	Teddy			Cones			APBP
	Nonocc	All	Disc	Nonocc	All	Disc	(%)
Proposed	**4.91**	**10.20**	**13.31**	**2.45**	**8.39**	**7.11**	**4.78**
AdaptiveGF	5.71	11.3	14.3	2.44	8.22	7.05	4.98
HistAggr2	5.88	11.3	14.7	2.41	8.18	7.21	5.20
CrossLMF	5.50	10.6	14.2	2.34	7.82	6.80	5.13

Table 1. Bad pixel error of the proposed stereo matching algorithm based on the Middlebury stereo evaluation benchmark version 2, and comparison with existing local stereo matching methods.

matching, a novel sparse disparity subset selection method is also proposed by assigning a significance status to candidate disparity values, and selecting the significant disparity values adaptively. For this purpose, a fast cost aggregation is conducted first to generate an initial coarse aggregated cost volume for all pixels at all disparity levels, and then a sparse disparity subset is estimated from the initial coarse aggergated cost volume for each pixel by adaptively selecting both the local minima and neighborhood points near local minima of initial aggregated matching cost. An adaptive guided filtering method is proposed based on the estimate sparse disparity subset to generate the refined aggregated cost volume. Finally, the disparity maps are calculated using a winner-take-all method and are refined through occlusion handling and post-processing steps [4].

The proposed stereo matching algorithm is tested on Middlebury stereo evaluation benchmark images. The bad pixel error of the resulting disparity maps for different datasets are shown in Table 1. The experimental results show that the proposed method produces disparity results with lower disparity bad pixel errors compared with existing local stereo matching methods [6, 10, 15]. Further results also show that using the estimated sparse disparity subset for cost aggregation performs better than using the full disparity range.

3. Conclusion

In this paper, we we propose a novel Sparse Locally Adaptive Cost aggregation (SLAC) local stereo matching method. We propose a novel cost aggregation method performed in a locally adaptive manner by adapting the support region to the local image intensity and structure, using the robustly estimated disparity subset and adaptive guided filter weights. The experimental results show that the proposed method outperforms most recent local stereo matching algorithms [6, 10, 15], and is capable to achieve state-of-the-art disparity results while preserving linear computation efficiency.

References

[1] A. Buades and G. Facciolo. Reliable multiscale and multiwindow stereo matching. *SIAM Journal on Imaging Sciences*, 8(2):888–915, 2015.

[2] Anonymous. The fast bilateral solver. *ECCV 2016 submission 527*, 2016.

[3] Anonymous. Image-guided non-local dense matching with three-steps optimization. *ISPRS Congress 2016 submission 231*, 2016.

[4] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual correspondence and beyond. *2011 IEEE CVPR*, pages 3017–3024, 2011.

[5] C. Zhang, Z. Li, Y. Cheng, R. Cai, H. Chao, and Y. Rui. Meshstereo: A global stereo model with mesh alignment regularization for view interpolation. *2015 IEEE CVPR*, pages 2057–2065, 2015.

[6] D. Min, J. Lu, and M. N. Do. Joint histogram based cost aggregation for stereo matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):2539–2545, 2013.

[7] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frames stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, 2002.

[8] F. Guney and A. Geiger. Displets: Resolving stereo ambiguities using object knowledge. *2015 IEEE CVPR*, pages 4165–4175, 2015.

[9] J. Jiao, R. Wang, W. Wang, S. Dong, Z. Wang, and W. Gao. Local stereo matching with improved matching cost and disparity refinement. *IEEE Multimedia*, 21(4):16–27, 2014.

[10] J. Lu, K. Shi, D. Min, L. Lin, and M. N. Do. Cross-based local multipoint filtering. *2012 IEEE International Conference on Computer Vision and Pattern Recognition*, pages 430–437, 2012.

[11] J. Zbontar and Y. LeCun. Computing the stereo matching cost with a convolutional neural network. *2015 IEEE CVPR*, pages 1592–1599, 2015.

[12] K. J. Yoon and I. S. Kweon. Locally adaptive support-weight approach for visual correspondence search. *2005 IEEE CVPR*, 2:924–931, 2005.

[13] K. Zhang, J. Lu, and G. Lafruit. Cross-based local stereo matching using orthogonal integral images. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(7):1073–1079, 2009.

[14] M. A. Helala and F. Z. Qureshi. Accelerating cost volume filtering using salient subvolumes and robust occlusion handling. *Computer Vision–ACCV 2014*, pages 316–331, 2014.

[15] Q. Yang, P. Ji, D. Li, S. Yao, and M. Zhang. Fast stereo matching using adaptive guided filtering. *Image and Vision Computing*, 32(3):202–211, 2014.

[16] Y. Peng, G. Li, R. Wang, and W. Wang. Stereo matching with space-constrained cost aggregation and segmentation-based disparity refinement. *International Society for Optics and Photonics in Electronic Imaging*, pages 939309–939309, 2015.